

SCIENTIFIC REPORTS



OPEN

High-throughput sequencing analysis revealed the regulation patterns of small RNAs on the development of *A. comosus* var. *bracteatus* leaves

Ying-Yuan Xiong¹, Jun Ma¹, Ye-Hua He², Zhen Lin¹, Xia Li¹, San-Miao Yu¹, Rui-Xue Li¹, Fu-Xing Jiang¹, Xi Li¹, Zhuo Huang¹ & Ling-Xia Sun¹

Studies of the molecular mechanisms involved in the formation of the albino leaf cells are important for understanding the development of chimera leaves in *Ananas comosus* var. *bracteatus*. In this study, we identified a total of 163 novel miRNAs involved in the development of complete white (CWh) and complete green (CGr) leaves using high-throughput sequencing method. The potential miRNA target genes were predicted and annotated using the NR, Swiss-Prot, GO, COG, KEGG, KOG and Pfam databases. The main biological processes regulated by miRNAs were revealed. The miRNAs which potentially play important roles in the development of the leaves and the albino of the CWh leaf cells were selected and their expression patterns were analyzed. The expression levels of nine miRNAs and their potential target genes were studied using qRT-PCR. These results will help to elucidate the functional and regulatory roles of miRNAs in the formation of the albino cells and the development of the leaves of *A. comosus* var. *bracteatus*. These data may also be helpful as a resource for studies of small RNA in other leaf color chimeric plant species.

Ananas comosus var. *bracteatus* (red pineapple) is a commercially cultivated monocot plant originating from South America. It belongs to the *Bromeliaceae* family. While other plants in this family are valued for their fruit, the high quality silk fiber of the stem and leaves^{1,2}, and a large number of secondary metabolites^{3–8}, the red pineapple is cultivated commercially as an ornamental plant because of its colorful chimera leaves. The chimera leaves consist of normal green cells and albino white cells, but the chimeric character of the leaves in cultivated plants is particularly unstable. The leaf color of the regenerated plants varies significantly during tissue culture, in which only about 1% of the regenerated plants have the same chimeric character as the mother plant. Most of the regenerated plants lose their economic value due to the lack of colorful chimera leaves. In order to understand the molecular mechanism of the formation of the albino cells and the development of the leaves, we undertook transcriptome sequencing of the CGr cells and the CWh cells. The results of transcriptome sequence revealed the significant differences in transcription levels between the CGr and CWh plants, especially in those aspects that control photosynthesis, porphyrin and chlorophyll metabolism, and carotenoid biosynthesis⁹. These studies provided a frame of transcription variations between CGr and CWh plants that most likely result in the formation of the albino leaf cells. Based on these results, we could further study the regulatory roles of small RNA presumably involved in the development of albino leaves.

MicroRNAs are the short non-coding RNAs (20–24 nucleotides) that silence the expression of target genes through complementary sequence matching¹⁰. In plants, miRNAs control the expression of various genes involved in the processes of development, growth, and physiology. In particular, these include genes that encode transcription factors, stress response proteins, and other proteins¹¹. Although the regulation of genes by miRNA is indirect, it is believed that miRNAs play important roles in plant growth and development¹².

¹College of Landscape Architecture of Sichuan Agricultural University, Chengdu, Sichuan, 611100, China.

²Horticultural Biotechnology College of South China Agricultural University, Guangzhou, Guangdong, 510642, China. Correspondence and requests for materials should be addressed to J.M. (email: junma365@hotmail.com)

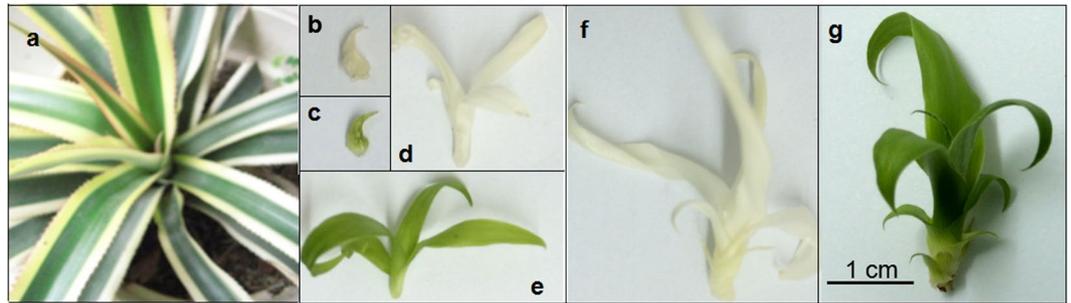


Figure 1. Materials for miRNA sequence. (a) Wild chimeric plant of *A. comosus* var. *Bracteatus*. (b) Complete white shoots at developmental stage 1 (with unexpanded leaves, WS1). (c) Complete green shoots at developmental stage 1 (GS1). (d) Complete white shoots at developmental stage 2 (with four to five expanded leaves, WS2). (e) Complete green shoots at developmental stage 2 (GS2). (f) Complete white shoots at developmental stage 3 (with ten to twelve expanded leaves, WS3). (g) Complete green shoots at developmental stage 3 (GS3).

In this study, small RNA libraries were constructed at the different developmental stages of CWh and CGr leaves. 163 novel miRNAs and their potential target genes were identified by Illumina sequencing. To our knowledge, this is the first comparative miRNA high-throughput sequencing of *A. comosus* var. *bracteatus*. This study was mainly focused on revealing the expression profiles and regulation patterns of miRNAs involved in the CGr and CWh leaf development. The miRNAs which play potentially important roles in the development of the leaves and the albino of the CWh leaf cells were selected and analyzed. The expression levels of nine miRNAs and their potential target genes were detected by real-time qPCR. Our results provide a valuable resource for studying microRNA regulation involved in the leaf color formation and leaf development of *A. comosus* var. *bracteatus*.

Materials and Methods

Plant material. Complete green (GS1, GS2, GS3) and white (WS1, WS2, WS3) shoots were derived from the stem explants of chimera plants of *A. comosus* var. *bracteatus* (Fig. 1a) by tissue culture, in accordance with our previous protocol. Three developmental stages were evaluated based on the number of expanded leaves on shoots. WS1 (Fig. 1b) and GS1 (Fig. 1c) are complete white and green shoots respectively with only unexpanded leaves; WS2 (Fig. 1d) and GS2 (Fig. 1e) are complete white and green shoots respectively with four to five expanded leaves; WS3 (Fig. 1f) and GS3 (Fig. 1g) are complete white and green shoots respectively with ten to twelve expanded leaves. The leaves were collected at the above three stages, immediately frozen in liquid nitrogen, and then stored at -80°C until analysis.

Measurement of chlorophyll and carotenoid contents. Complete green (GS1, GS2, GS3) and complete white (WS1, WS2, WS3) leaves were selected for chlorophyll and carotenoid measurements. The contents of chlorophyll and carotenoid concentrations were measured using the Holm equation and the method as previously described¹³.

RNA extraction, small RNA library construction and sequencing. Total RNA was isolated from six samples using Trizol (Invitrogen, USA) according to the manufacturer's instructions. The purity, concentration and integrity of RNA samples was analyzed using Nanodrop, Qubit 2.0, the Agilent 2100 bioanalyzer respectively in order to ensure the high quality of sample for sequencing. A total amount of $1.5\ \mu\text{g}$ RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEB Next Ultra small RNA Sample Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v4-cBot-HS (Illumina) according to the manufacturer's instructions. After the cluster generation, the library preparations were sequenced on an Illumina HiSeq 2500 platform and single-end 50 nt reads were generated (Beijing Biomarker Technologies Co., Ltd, Beijing, China).

Identification and bioinformatics analysis of miRNAs. Raw data (raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads containing ploy-N and low quality reads from raw data. And reads were trimmed and cleaned by removing the sequences smaller than 18 nt or longer than 30 nt. Using Bowtie tools soft, the Clean Reads were aligned respectively with Silva database, GtRNadb database, Rfam database and Rfam database to filter ribosomal RNA (rRNA), transfer RNA (tRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and other ncRNA and repeats. The clean reads were mapped to genomic sequences (Acomosus_321_v3, https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Acomosus_er). The miRDeep2¹⁴ was used to detect the known miRNA and new miRNA by comparing with known miRNAs from miRBase.

Quantification of miRNA expression levels and identification of differentially expressed miRNA. miRNA expression levels were estimated for each sample: sRNA were mapped back onto the precursor sequence. Read count for each miRNA was obtained from the mapping results. The differentially expressed miRNA was

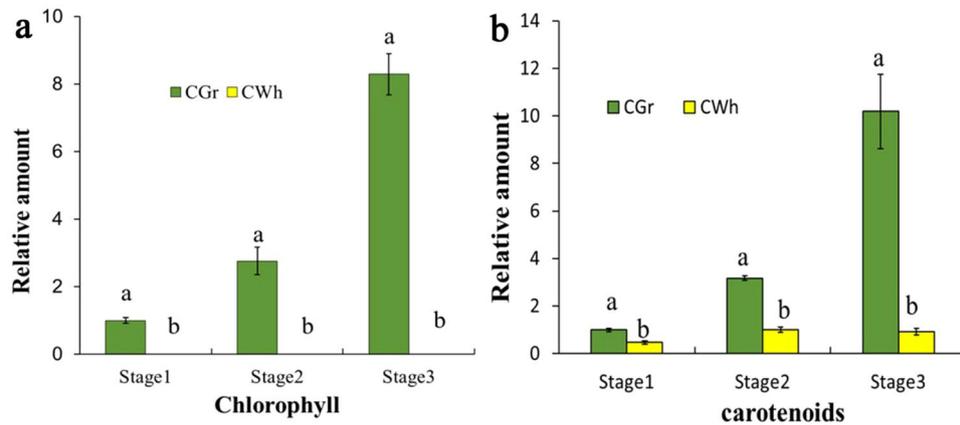


Figure 2. Pigment concentrations of the CGr and CWh leaves. **(a)** The concentration of chlorophyll in the CGr and CWh leaves at three developmental stages. **(b)** The concentration of carotenoids in the CGr and CWh leaves at three developmental stages. The relative values of pigments concentrations were calculated using the value of CGr leaves at developmental stage 1 as 1. Different letters in columns indicate statistically significant differences ($P < 0.01$) according to a T-test.

identified by IDEG6¹⁵. The rules used for identification of differentially expressed miRNA were $|\log_2(\text{FC})| > = 1$ and $\text{FDR} < = 0.01$.

Prediction of miRNA target genes and function annotation. The potential target genes of sequenced miRNAs were predicted by Target Finder¹⁶. Gene function was annotated based on the following databases: Nr (NCBI non-redundant protein sequences)¹⁷; Nt (NCBI non-redundant nucleotide sequences); Pfam (Protein family)¹⁸; KOG/COG (Clusters of Orthologous Groups of proteins)^{19,20}; Swiss-Prot (A manually annotated and reviewed protein sequence database)²¹; KO (KEGG Ortholog database); GO (Gene Ontology)²² and KEGG (<http://www.genome.jp/kegg/kaas/>) database.

qRT-PCR identification of the expression levels of miRNAs and potential target genes. Total RNAs of six samples as used in transcriptome sequence were reverse transcribed into cDNA using the TURE script 1st Stand cDNA SYNTHESIS Kit (Aidlab) with miRNA specific primer (Table S1). qRT-PCR was performed using the analytikjena-qTOWER2.2 Real-Time PCR System (Germany) with SYBR[®] Green (DBI). PCR reactions of miRNAs were performed as follow: 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s, and 58 °C for 30 s. *U6* gene was used as internal control. PCR reactions of potential target genes were performed at 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 58 °C for 30 s. A melting curve (at 60–95 °C) was generated to check amplification specificity. Each reaction presented three biological repeats. The primers used were listed in Table S2.

Results

Chlorophyll and carotenoid content analysis of the CWh and CGr leaves among three developmental stages. The content of chlorophyll in CWh leaves is not detectable among all three developmental stages, while that in CWh leaves is significantly high, and increases gradually with the development of plants. (Fig. 2). These results confirmed that the CWh and CGr leaves were the typical presentations of the albino white cells and the normal green cells, respectively. The content of carotenoid of the CGr leaves is greatly higher than that in the CWh leaves among all three developmental stages, and it is gained gradually in both CGr and CWh leaves along the development of the plants (Fig. 2).

High-throughput small RNA sequencing of *Ananas comosus* var. *bracteatus* leaves. Our former studies had confirmed that chlorophyll in CWh leaves is significantly suppressed, and that the transcriptional pattern of CWh leaves is different from that of CGr leaves⁹. In order to identify miRNAs and to reveal the regulation mechanisms of gene expression in the *A. comosus* var. *bracteatus* leaves, the high-throughput sequencing of six miRNA libraries constructed from the CGr and CWh leaves at three developmental stages were performed. After filtering the low-quality reads, we obtained 16–18 million clean reads from each sample (Table S3). Among them, 5.8–7.4 million (about 35–45%) reads were perfectly mapped to the *A. comosus*_321_v3 genome. All of the clean reads mapped to the *A. comosus*_321_v3 genome were used for miRNA identification by miRDeep2. To facilitate the access and use of the sequence data, the data have been submitted to NCBI with BioProject number: PRJNA389361.

The different types of total small RNA sequences were further classified after using the Bowtie software to blast the clean reads against the Silva, GtRNadb, Rfam and Rепbase databases²³. The clean reads were annotated as tRNAs, rRNA, snRNA, scrRNA, snoRNAs, and repeat-associated small RNAs based on the Rfam database (Table S4).

Identification of miRNAs in *A. comosus* var. *Bracteatus*. All small RNA sequences were aligned to the precursor/mature miRNAs in the miRbase database using miRDeep2 software²⁴. No known-miRNA was predicted. A total of 162, 163, 162, 162, 163, 161 novel miRNAs were identified in the six samples, respectively. The sequences of the 163 miRNAs are listed in Table S5. Among the 163 miRNAs, 104 miRNAs are conservative and 59 miRNAs are un-conservative. The length distribution of the identified miRNAs was shown in Fig. S1a. The most abundant sequence size is 21 nt-long, followed by 24 nt. The base bias on the first site of miRNAs with specific lengths and a specific site of miRNAs were shown in Fig. S1b,c. The majority of the 163 miRNAs started with a 5'-U, except for the 24nt-long miRNAs, which mostly started with A (59.26%). All the 19 nt-long miRNAs started with a 5'-U, followed by 21 nt-long miRNA (70.31%). The (A + U) percentage of the first nucleotide bias of 21 nt-long and 24-long miRNAs is 85.94% and 75.93% respectively. Nucleotide bias analysis at each position indicated that (A + U) occupied a very high percentage (77.16%) at the start of reads, while (G + C) occupied high frequency (61.02% at 24 nt and 80% at 25 nt) at the ends of reads. The distribution of the 163 identified miRNAs in each miRNA family was shown in Fig. S2. The 163 miRNAs were predicted to belong to 46 miRNA families. The most enriched family is miR399 (10), followed by miR159 (6), miR160 (6), miR164 (6) and miR166 (5).

Prediction and annotation of miRNA putative target genes. The putative target genes of miRNA were predicted by Target Finder software. Among the 163 miRNAs, 123 miRNAs were predicted with 488 potential target genes. The potential target genes were blasted to the Nr, Nt, Swiss-Prot, GO, COG, KEGG, KOG and Pfam databases. The statistic of the functional annotation of the target genes were listed in Table S6.

The Nr homologous species distribution of the 414 candidate target genes were shown in Fig. S3a. The blasted species are all monocots. The first three enriched species are *Elaeis guineensis* (166), *Phoenix dactylifera* (123) and *Musa acuminata* (42), which are all perennial tropical monocots. The blast results of the potential target genes found in COG database were shown in Fig. S3b. A total of 155 target genes were assigned to COG classification. The target genes were classified into 25 categories. The largest group is 'General function prediction only', followed by 'Transcription' and 'Replication, recombination and repair', and then 'Signal transduction mechanisms'. GO analysis was conducted to describe the functional classification of target genes in terms of their associated biological processes, cellular components, and molecular functions. A total of 243 target genes were annotated by GO database (Fig. S3c). 'Cell part', 'Cell' and 'Organelle' are the most enriched groups under the cellular component category. The most highly represented groups of molecular function category are 'Binding' and 'Catalytic activity'. For the biological process category, the 'Metabolic process', 'Cellular process' and 'Single-organism process' are the most enriched groups. The potential target genes were also blasted to the KEGG database using BLASTx. Forty-two KEGG pathways were annotated by the target genes (Fig. S3d). The most frequently represented pathways are 'Biosynthesis of amino acids' (ko01230, 5), 'Oxidative phosphorylation' (ko00190, 5), followed by 'Tropane, piperidine and pyridine alkaloid biosynthesis' (ko00960, 4) and 'mRNA surveillance pathway' (ko03015, 4).

Expression profiles of the miRNAs in different samples. The expression levels of the 163 miRNAs in the six samples were digitally measured and their expression relationships were characterized. The expression levels of the miRNAs were analyzed through the heat map method (Fig. 3a). The six samples were clustered to three groups, GS1 and WS1, WS2 and WS3, GS2 and GS3. Ten clusters were obtained from the expression data of the six samples using K-means method (Fig. 3b). These miRNAs target on genes functioned in transcription and translation, pigment biosynthesis and chloroplast development, sulfate assimilation, starch synthase, disease resistance ect. These miRNAs may presumably play very important roles in the growth and development of the complete white leaves.

The statistics of the number of differentially expressed miRNAs was shown in Fig. 3c. There are total 91 and 80 differentially expressed miRNAs between stage 1 and stage 2, in the CGr and CWh leaves respectively and about half of the differentially expressed miRNAs is down-regulated. Moreover, there are total 53 and 69 differentially expressed miRNAs between stage 2 and 3 in the CGr and CWh leaves respectively, and 87% and 84% of the differentially expressed miRNAs are down-regulated between stage 2 and 3 in CGr and CWh leaves respectively. This indicated that the number of differentially expressed miRNAs is decreased along the development of the plants in both CGr and CWh leaves. Compared with CGr and CWh leaves, there are 29, 65, 69 differentially expressed miRNAs at stage 1, 2, 3 respectively. At stage 2 and stage 3, the numbers of up-regulated (41, 40) and down-regulated miRNAs (24, 29) among CGr and CWh leaves are similar.

In order to identify potentially important miRNAs which functioned in the albino and development of the CWh leaves, the differentially expressed miRNAs ($|\log_2(\text{FC})| \geq 3$ and $\text{FDR} \leq 0.01$) between CGr and CWh leaves are selected and analyzed (Table 1). The expression of all the miRNAs except *Ab-miR131*, *Ab-miR132*, *Ab-miR139* and *Ab-miR152* are down regulated in CWh leaves. It is worth mentioning that the expression values of the differentially expressed miRNAs between CGr and CWh at three developmental stages are different, except *Ab-miR7*. The $\log_2\text{FC}$ (WS/GS) value of *Ab-miR7* is -3.84 and -5.83 at developmental stages 2 and 3, respectively. *Ab-miR7* is predicted to target on genes play roles in photosynthesis, sulfur metabolism, non-homologous end-joining and translation initiation.

Expression patterns of miRNAs targeted genes related to transcription factors. The expression patterns of the miRNAs target on transcription factors were shown in Table 2. There are fifteen miRNAs target on twelve transcription factors. Among them, *Ab-miR152* and *Ab-miR39* are differentially expressed between CGr and CWh leaves at all three developmental stages. *Ab-miR87* and *Ab-miR124* are differentially expressed between CGr and CWh leaves at only developmental stage 1 and 2.

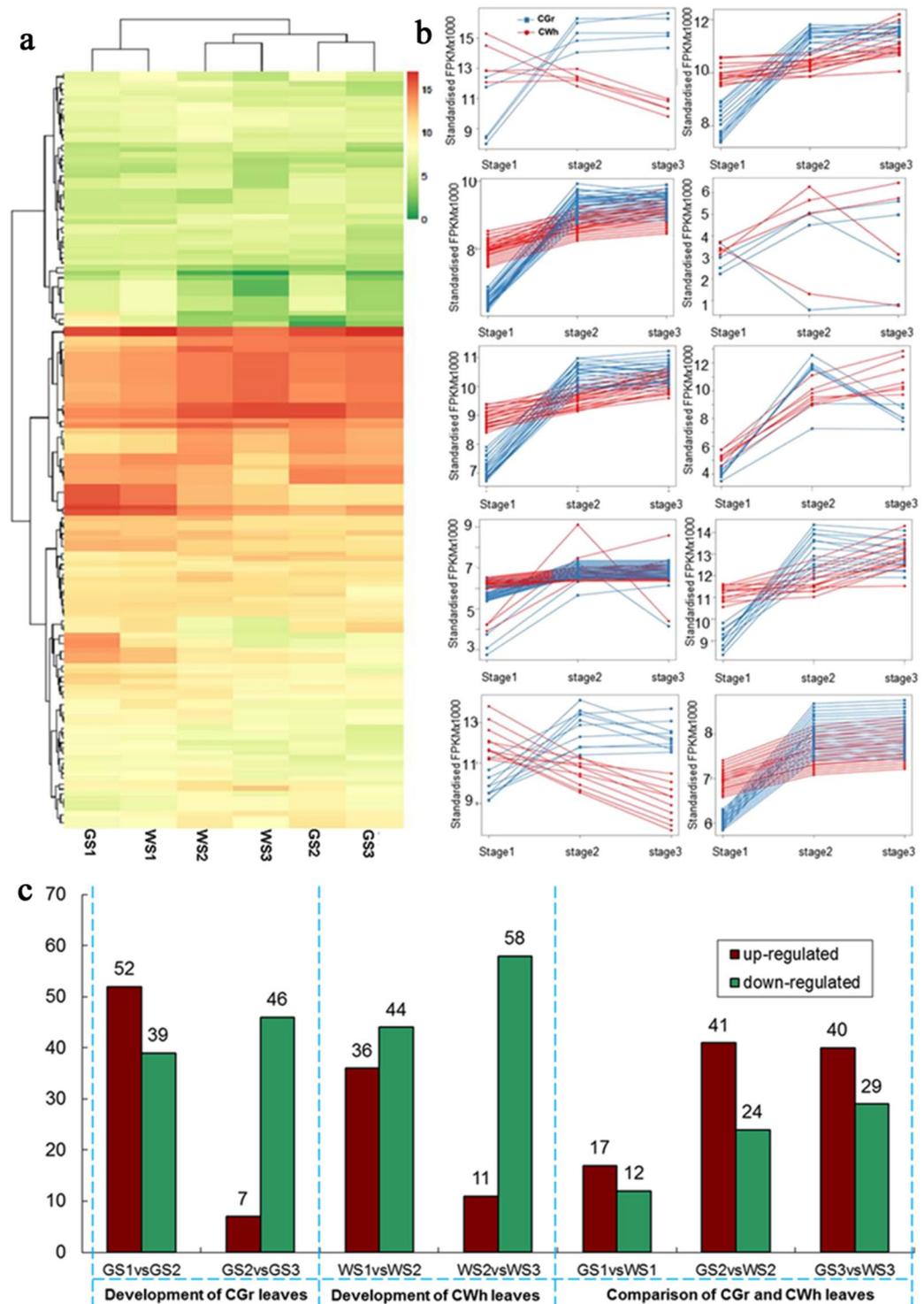


Figure 3. Expression and clustering analysis of the miRNAs. (a) Heat map of differentially expressed miRNAs detected in this sequence based on their expression in 6 samples. (b) Cluster analysis of miRNAs with the K-means method. (c) Comparison of the number of differentially expressed miRNAs in the six samples. *Development of CGr leaves*: number of differentially expressed miRNAs in CGr leaves at three developmental stages. *Development of CWh leaves*: number of differentially expressed miRNAs in CWh leaves at three developmental stages. *Comparison of CGr and CWh leaves*: number of differentially expressed miRNAs between CGr and CWh leaves at three developmental stages. *Up-regulated*: the expression level of the miRNA is up-regulated in the latter sample compare to the former sample in two compared samples. *Down-regulated*: the expression level of the miRNA is down-regulated in the latter sample compare to the former sample between two compared samples.

Stage	MiRNA	Family	Target gene functions	log2FC(WS/GS)
GS1vsWS1	<i>Ab-miR45</i>	MIR1319	SNARE interactions in vesicular transport Oxidative phosphorylation	-3.23
	<i>Ab-miR84</i>		Endocytosis	-4.08
	<i>Ab-miR106</i>	MIR482	Transcription Protein phosphorylation Development and cell death	-3.33
	<i>Ab-miR107</i>	MIR482	Same as <i>Ab-miR106</i>	-3.33
	<i>Ab-miR124</i>		Biosynthesis of amino acids Photosynthesis Response to stress	-3.44
GS2vsWS2	<i>Ab-miR6</i>	MIR395	Photosynthesis Sulfur metabolism Non-homologous end-joining Translational initiation	-3.03
	<i>Ab-miR7</i>	MIR395	Same as <i>Ab-miR6</i>	-3.48
	<i>Ab-miR46</i>	MIR399	Plant hormone signal transduction	-3.72
	<i>Ab-miR152</i>	MIR1120	Transcription factor bHLH	3.19
	<i>Ab-miR144</i>	MIR399	rRNA methylation	-3.45
	<i>Ab-miR139</i>	MIR1516	No target gene predicted	3.31
GS3vsWS3	<i>Ab-miR7</i>	MIR395	Same as <i>Ab-miR6</i>	-5.83
	<i>Ab-miR8</i>	MIR395	Sulfur metabolism Signal transduction Leaf development Translational initiation Non-homologous end-joining	-4.66
	<i>Ab-miR11</i>	MIR528	Oxidation-reduction process Starch synthase Phosphorylation Endocytosis	-3.19
	<i>Ab-miR27</i>	MIR528	Same as <i>Ab-miR11</i>	-3.25
	<i>Ab-miR101</i>	MIR528	Same as <i>Ab-miR11</i>	-3.25
	<i>Ab-miR131</i>	MIR171_2	No target gene predicted	3.24
	<i>Ab-miR132</i>		No target gene predicted	3.62

Table 1. Differentially expressed miRNAs ($|\log_2(\text{FC})| \geq 3$ and $\text{FDR} \leq 0.01$) between CGr and CW h leaves.

Identification of miRNAs target on genes related to chlorophyll biosynthesis and photosynthesis.

The differentially expressed miRNAs with target genes related to chlorophyll biosynthesis and photosynthesis are listed in Table 3. There are seven miRNAs predicted to target on chlorophyll biosynthesis and nine miRNAs targeted on photosynthesis. *Ab-miR73*, *Ab-miR90* and *Ab-miR98* are predicted to target on Ferrochelatase (HemH, EC: 4.99.1.1), and are down regulated in CGr leaves at developmental stage 2. *Ab-miR127*, *Ab-miR160*, *Ab-miR161* and *Ab-miR163* are targeted on divinyl chlorophyllide a 8-vinyl-reductase (DVR, EC: 1.3.1.75), which are down regulated at stage 1 and up regulated at stages 2 and 3 in CW h leaves. And these four miRNAs are down regulated at developmental stage 2 compare to developmental stage 1 in both CGr and CW h leaves. *Ab-miR6*, *Ab-miR7* and *Ab-miR8* targeted on PGR-5 like protein (Ferredoxin-plastoquinone reductase 1), which is involved in the cyclic electron flow (CEF) around photosystem I²⁵. *Ab-miR6*, *Ab-miR7* and *Ab-miR8* are up regulated at stage 1 and down regulated at stage 2 and stage 3 in CW h leaves. *Ab-miR124* is predicted to target on trihelix transcription factor, which may act as a molecular switch in response to light signals²⁶, and which is significantly down regulated in CW h leaves at developmental stage 1 and 2. *Ab-miR124* is significantly down regulated along the development of both CGr and CW h leaves. The other miRNAs listed are targeted on genes related to the reaction center of the photosynthesis system^{27–29}.

Expression profiles of nine miRNAs and their potential target genes. In order to reveal the potential regulatory mechanism of the miRNAs action on the target genes, the expression profiles of nine miRNAs presumably relating to photosynthesis and chlorophyll biosynthesis, and their potential target genes were detected using qRT-PCR. The expression patterns of these miRNAs detected by RNA-seq are well corroborated by qRT-PCR (Fig. 4). This confirmed that the sequence results are reliable. The expression levels of the nine potential target genes of the nine miRNAs analyzed by qRT-PCR are shown in Fig. 5. Our results showed that the expression levels of these target genes are not all dependent on their predicted miRNAs based on the analysis of their expression levels. Many factors may affect the accuracy of target gene prediction and the regulatory relationships between miRNA and their target genes. For instance, the lacking of relevant research and genomic information may affect the accuracy of target gene prediction³⁰. It is the first report of miRNA sequencing in *A. comosus* var. *bracteatus*, and the identification of miRNA is based on the genomic information of *A. comosus* var. *comosus*. Furthermore, some of these potential target genes are predicted to be regulated by more than one miRNAs at the translational level³⁰. It is reported that isomeric sequences shared the authentic miRNA target genes with canonical miRNA³¹.

DEG ID	Target gene annotation	GS1vsWS1	GS2vsWS2	GS3vsWS3	GS1vsGS2	GS2vsGS3	WS1vsWS2	WS2vsWS3
<i>Ab</i> -miR73 <i>Ab</i> -miR90	AP2-like ethylene-responsive transcription factor	—	-2.95	—	1.84	—	—	1.49
<i>Ab</i> -miR37, <i>Ab</i> -miR152, <i>Ab</i> -miR52	bHLH49 isoform X1 bHLH18-like bHLH140-like	— 3.19 —	— 2.26 —	1.06 -2.22 —	— — —	— — -1.34	— 1.09 —	— -1.86 -1.48
<i>Ab</i> -miR37	GAMYB-like isoform X2	—	—	1.06	—	—	—	—
<i>Ab</i> -miR68	homeobox-leucine zipper protein	—	—	1.20	2.46	-1.23	1.68	—
<i>Ab</i> -miR52, <i>Ab</i> -miR87	WRKY48 WRKY9	— 1.82	— -2.32	— —	— 1.29	-1.34 -4.10	1.09 -2.84	-1.48 -1.74
<i>Ab</i> -miR9	CCAAT-binding transcription factor	—	—	-1.62	-1.92	—	-1.62	—
<i>Ab</i> -miR18	TCP family transcription factor	—	1.46	—	-3.83	—	-2.04	—
<i>Ab</i> -miR94	LIM domain-containing protein WLIM1	—	—	2.32	1.89	-2.17	—	1.04
<i>Ab</i> -miR49	Transcription factor BRX N-terminal domain	—	1.16	—	—	—	—	-1.60
<i>Ab</i> -miR39	SRF-type transcription factor	1.60	-2.25	-1.09	1.49	-2.75	-2.36	-1.59
<i>Ab</i> -miR124	Trihelix transcription factor GT-3b	-3.44	-2.37	—	-6.09	—	-5.02	—
<i>Ab</i> -miR156	MADS-box transcription factor 50	—	2.13	—	-1.65	—	—	-1.91

Table 2. The expression patterns of the miRNAs targeted genes related to transcription factors. Note: the values in each column (AvsB) was the value of $[\log_2FC (B/A)]$. — means no significant differences were detected between the two samples.

DEG ID	Target gene annotation	GS1vsWS1	GS2vsWS2	GS3vsWS3	GS1vsGS2	GS2vsGS3	WS1vsWS2	WS2vsWS3
<i>Ab</i> -miR127 <i>Ab</i> -miR160 <i>Ab</i> -miR161 <i>Ab</i> -miR163	divinyl chlorophyllide a 8-vinyl-reductase	-1.10 -1.10 -1.10 -1.10	2.06 2.06 2.06 2.06	1.25 1.25 1.25 1.25	-5.32 -5.32 -5.32 -5.32	— — — —	-2.16 -2.16 -2.16 -2.16	— — — —
<i>Ab</i> -miR73 <i>Ab</i> -miR90 <i>Ab</i> -miR98	Ferrochelatase	—	-2.95 -2.95 -2.95	—	1.84 1.84 1.84	— — —	— — —	1.49 1.49 1.49
<i>Ab</i> -miR6 <i>Ab</i> -miR7 <i>Ab</i> -miR8	PGR5-like protein	1.35 1.32 1.39	-3.03 -3.48 -2.17	-5.50 -5.83 -4.66	— 1.03 —	1.04 1.15 1.02	-3.58 -3.76 -3.64	-1.42 -1.20 -1.47
<i>Ab</i> -miR124	Trihelix transcription factor	-3.44	-2.37	—	-6.09	—	-5.02	—
<i>Ab</i> -miR52 <i>Ab</i> -miR83	<i>PsbB</i> (CP47) tetrapyrrole-binding protein	— —	— —	— 1.60	— —	-1.34 -1.36	1.10 —	-1.48 —
<i>Ab</i> -miR108	Ferredoxin-thioredoxin reductase	—	1.16	—	—	—	—	-1.17
<i>Ab</i> -miR82 <i>Ab</i> -miR104	<i>PsbP</i> sedoheptulose-1,7-bisphosphatase	— —	— —	1.65 1.65	-3.19 -3.19	-1.29 -1.29	-2.08 -2.08	— —
<i>Ab</i> -miR3	Chlorophyll a-b binding protein	—	—	1.51	-1.55	-1.19	—	—
<i>Ab</i> -miR101	polyphenol oxidase	—	-1.66	-3.19	1.11	—	—	-1.77
<i>Ab</i> -miR11	polyphenol oxidase granule-bound starch synthase 1	—	-1.66 -1.67	-3.19 -3.25	1.11 1.10	— —	— —	-1.77 -1.81

Table 3. The expression levels of differentially expressed miRNAs with target genes related to chlorophyll biosynthesis and photosynthesis. Note: the values in each column (AvsB) was the value of $[\log_2FC (B/A)]$. — means no significant differences were detected between the two samples.

Discussion

The unstable chimeric character of the leaves in *A. comosus* var. *bracteatus* during tissue culture inhibits the commercial production of clones through tissue culture. Therefore, elucidating the albino mechanism of leaf cells is significant on the regulation of the stability of the chimera character in this plant. It is well known that miRNAs play important regulatory roles in gene expression at the post-transcriptional level by repressing the gene translation³². Previous studies have confirmed that miRNAs play a role in leaf development in higher plants^{32–34}. It is reported that miRNAs are involved in regulation of many developmental processes, such as meristem identity, cell division and organ separation^{35–37}.

In this study, we identified miRNAs from CGr and CWh leaves at three development stages by high-throughput sequencing. A total of 163 miRNAs were identified. These miRNAs belong to 46 miRNA families. The top five enriched miRNA families are miR399, miR159, miR160, miR164 and miR166. miR399 is involved in functions which respond to nutrient deficiency^{38,39} and are also an important regulator of Pi starvation-induced root-associated acid phosphatase activity⁴⁰. miR159 negatively controls the expression of MYB33 and MYB65^{41,42}, and the expression of miR159 is suppressed by DELLAs⁴³. The miR159 works on leaf development and cell division by targeting TCP transcription factor genes^{44,45}. The miR160 targets the auxin response factors (ARFs) to

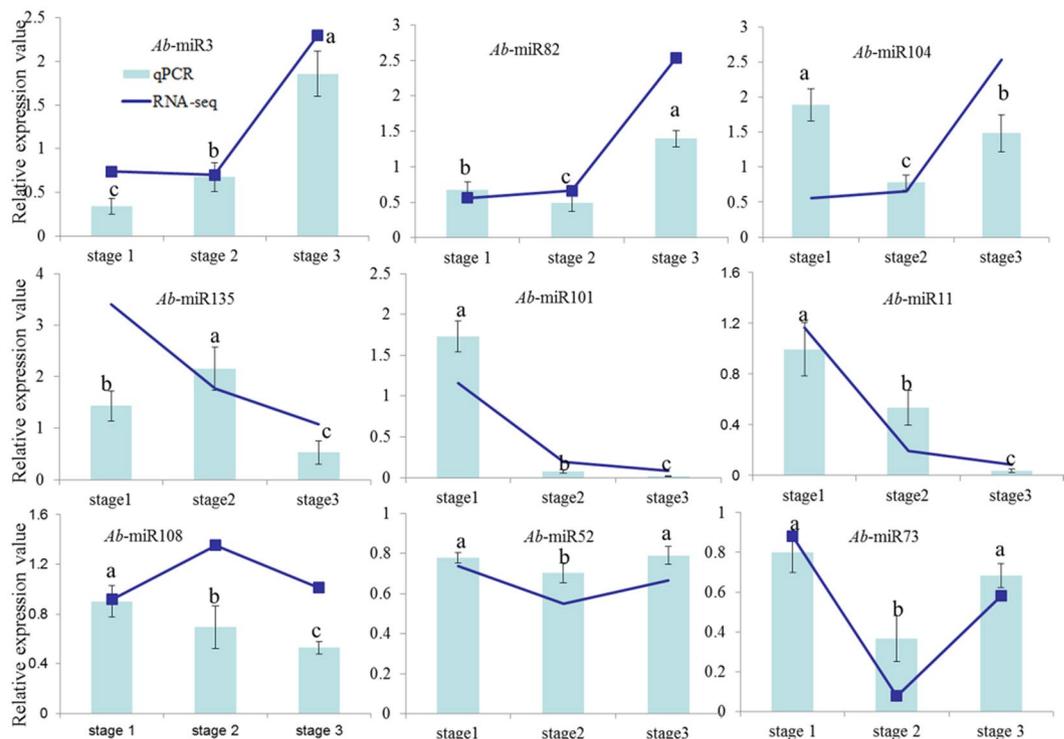


Figure 4. Expression profiles of miRNAs in the CGr and CWh leaves at three developmental stages. The relative expression values were the value of CWh/CGr. Each bar represents the mean value from triplicate experiments \pm SD. The letters marked on the columns indicate statistically significant differences ($P < 0.05$) according to a T-test.

negatively regulate the auxin signaling^{46–48}. The miR164 targets the NAC transcription factor ORE1⁴⁹ and CUC⁵⁰, both involved in functions related to leaf development. The miR 164 also functions in response to hormone signals by targeting the NAC genes^{50,51}. The miR166 regulates shoot apical meristem splitting by targeting the class III homeodomain leucine zipper transcription factors³². These results indicated that these growth and development aspects of *A. comosus* var. *bracteatus* may be regulated by miRNAs.

Along the development of both the CGr and CWh leaves, the number of differentially expressed miRNAs decreased. And the expression levels of about 85% of the differentially expressed miRNAs are decreased at development stage 3 compared to development stage 2. This data suggested that the regulation function of miRNA is more important at the early developmental stages of the leaves.

Since the expression levels of the seven miRNAs (*Ab-miR6*, *Ab-miR7*, *Ab-miR8*, *Ab-miR11*, *Ab-miR15*, *Ab-miR27*, *Ab-miR67*) are significantly decreased along the development of CWh leaves, while it is greatly increased in CGr leaves. These miRNAs are predicted to target on genes functioning in DNA repair, translation, sulfate assimilation, starch synthase, oxidation-reduction process, chloroplast development, leaf development, signal transduction etc. These results indicated that these biological processes are different between CGr and CWh leaves and regulated by miRNAs. The expression of most of the significantly differentially expressed miRNAs ($|\log_2(\text{FC})| \geq 3$ and $\text{FDR} \leq 0.01$) is down regulated in the CWh leaves. *Ab-miR39*, targeting on SRF-type transcription factor, is differentially expressed not only between CGr and CWh leaves at all three developmental stages, but also among the different developmental stages in the CGr and CWh leaves. SRF-type transcription factor plays roles on the positive regulation of transcription from RNA polymerase II promoter. *Ab-miR152* is differentially expressed between the CGr and CWh leaves along the development stages. *Ab-miR152* is a target on transcription factor bHLH18, which regulates transcription. *Ab-miR124* is significantly down regulated in CWh leaves and at the early development stage of leaves. It is predicted to target on Trihelix transcription factor GT-3b, which may play a role in the induction of CAM4 in response to pathogen and salt⁵³. It is predicted that these miRNAs could be presumably important to the albino and development of the leaves and are worth to further studying in the future.

The expression profiles of the nine miRNAs and their potential target genes in the CWh and CGr leaves showed that only a part of the potential target genes are dependent on their miRNAs. It has been previously shown that the expression profiles of all five detected potential target genes are independent of miRNAs in celery³². It is believed that the majority of miRNA regulation of genes is indirect⁵⁴ and not all target genes are directly regulated by miRNAs⁵⁵. Previous study has shown that the binding of miRNA is shown to up-regulate the expression of target mRNAs⁵⁶ in certain cellular conditions.

In this study, we reported the first comprehensive study on miRNAs of *A. comosus* var. *bracteatus* performed through high throughput sequencing and bioinformatics analysis methods. Our results showed that there are 163 novel miRNAs being identified, and their potential target genes are predicted and functionally annotated using

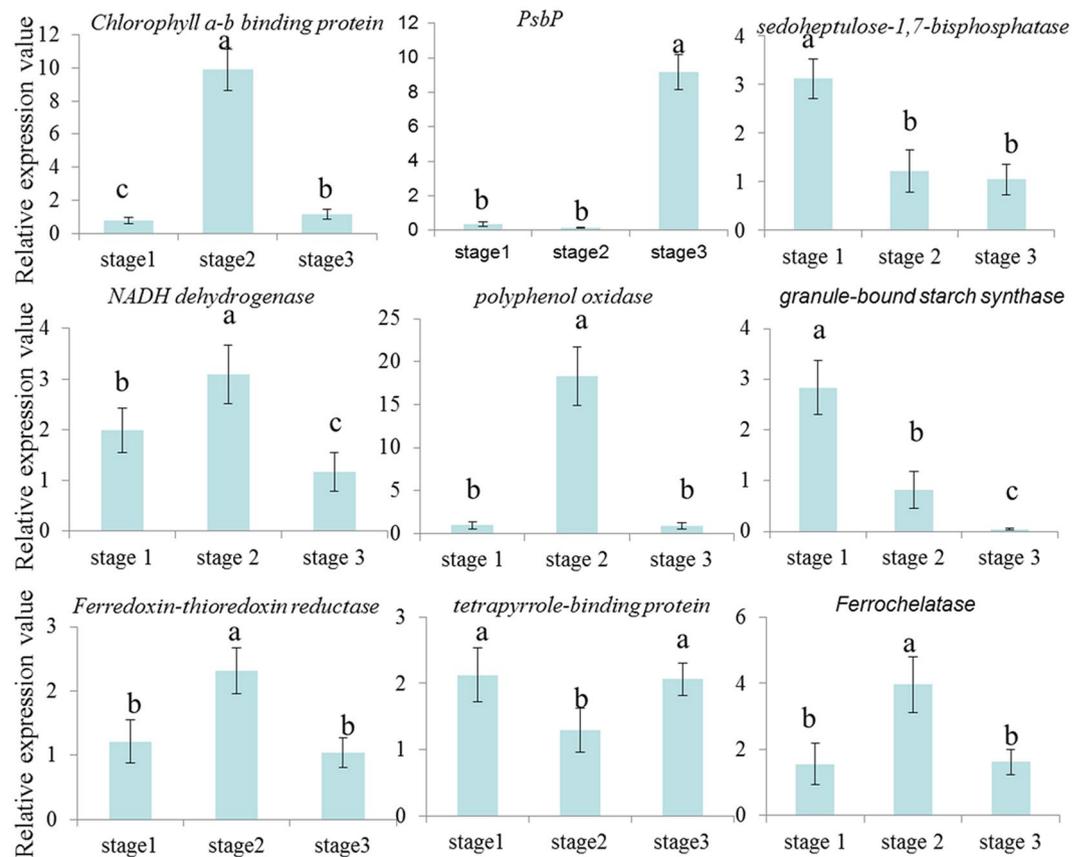


Figure 5. The expression patterns of the potential target genes at three developmental stages. The relative expression values were the value of CWh/CGr. Each bar represents the mean value from triplicate experiments \pm SD. The letters marked on the columns indicate statistically significant differences ($P < 0.05$) according to a T-test.

the Nr, Pfam, KOG/COG, Swiss-Prot, KO, GO and KEGG databases. The miRNAs presumably involved in the albino of the CWh leaves and the development of CGr and CWh leaves were selected and their expression patterns were analyzed. The expression of nine miRNAs related to chlorophyll biosynthesis and photosynthesis and their potential target genes were validated by qRT-PCR. These studies will provide a foundation on elucidating the functions and molecular regulatory mechanisms of miRNAs in the development of normal and albino leaves in *A. comosus* var. *bracteatus*.

References

- Collins, J. L. The Pineapple: Botany, Utilisation, Cultivation (Interscience Publishers, 1960).
- Montinola, L. R. P. (Amon Foundation, 1991).
- Wen, L., Wrolstad, R. E. & Hsu, V. L. Characterization of sinapyl derivatives in pineapple (*Ananas comosus* [L] Merr.) juice. *J Agric Food Chem.* **47**, 850–853 (1999).
- Takata, R. H. & Scheuer, P. J. Isolation of glyceryl esters of caffeic and p-coumaric acids from pineapple stems. *Lloyd.* **39**, 409–411 (1976).
- Taussig, S. J. & Batkin, S. Bromelain, the enzyme complex of pineapple (*Ananas comosus*) and its clinical application: an update. *J Ethnopharmacol.* **22**, 191–203 (1998).
- Rocha, F. D. & Kaplan, M. A. C. Secondary metabolites from *Ananas bracteatus* lindley (Bromeliaceae). *An Acad Bras Cienc.* **72**, 295 (2000).
- Bartholomew, D. P. The pineapple: botany, production, and uses (CABI, 2002).
- Beauman, F. The pineapple: king of fruits. 3674–3676 (Random House, 2006).
- Li, X. *et al.* Physiological characterization and comparative transcriptome analysis of white and green leaves of *Ananas comosus* var. *bracteatus*. *PLoS One.* **12**(1), 1–17 (2017).
- Cai, Y. *et al.* A brief review on the mechanisms of miRNA regulation. *Genomics Proteomics Bioinformatics.* **7**, 147–154 (2009).
- Rogersa, K. & Chen, X. M. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell.* **25**, 2383–2399 (2013).
- Shahab, S. *et al.* The effects of microRNA transfections on global patterns of gene expression in ovarian cancer cells are functionally coordinated. *BMC Medical Genomics.* **5**(1), 1–16 (2012).
- Holm, G. Chlorophyll mutation in barley. *Acta Agric Scandinavica.* **1**, 457–471 (1954).
- Zhang, Z., Jiang, L., Wang, J. J., Gu, P. Z. & Chen, M. An integrated tool for the identification of miRNA-target interaction in plants. *Bioinformatics.* **31**(2), 290–291 (2015).
- Romualdi, C., Bortoluzzi, S., D'Alessi, F. & Danieli, G. A. IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol Genomics.* **12**(2), 159–162 (2003).

16. Allen, E., Xie, Z., Gustafson, A. M. & Carrington, J. C. MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*. **121**, 207–221 (2005).
17. Deng, Y. Y. *et al.* Integrated nr database in protein annotation system and its localization. *Computer Engineering*. **32**(5), 71–74 (2006).
18. Eddy, S. R. Profile hidden markov models. *Bioinformatics Italic*. **14**(9), 755–763 (1998).
19. Koonin, E. V. *et al.* A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biology Italic*. **5**(2), 7 (2004).
20. Roman, L. T. *et al.* The COG database: a tool for genome scale analysis of protein functions and evolution. *Nucleic Acids Res*. **28**(1), 33–38 (2000).
21. Apweiler, R. *et al.* UniProt: the universal protein knowledge base. *Nucleic Acids Research*. **32**, 115 (2004).
22. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet*. **25**, 25–29 (2000).
23. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*. **10**(3), R25 (2009).
24. Zhang, Z. *et al.* MTide: an integrated tool for the identification of miRNA-target interaction in plants. *Bioinformatics*. **31**, 290–191 (2015).
25. DalCorso, G. *et al.* Complex containing PGRL1 and PGR5 is involved in the switch between linear and cyclic electron flow in *Arabidopsis*. *Cell*. **132**, 273–285 (2008).
26. Hiratsuka, K., Wu, X., Fukuzawa, H. & Chua, N. H. Molecular dissection of GT-1 from *Arabidopsis*. *Plant Cell*. **6**, 1805–1813 (1994).
27. Glauser, D. A., Bourquin, F., Manieri, W. & Schurmann, P. Characterization of ferredoxin: thioredoxin reductase modified by site-directed mutagenesis. *J Biol Chem*. **279**, 16662–16669 (2004).
28. Liu, J. *et al.* PsbP-domain protein1, a nuclear-encoded thylakoid lumenal protein, is essential for photosystem I assembly in *Arabidopsis*. *Plant Cell*. **24**, 4992–5006 (2012).
29. Adamiec, M. *et al.* Excitation energy transfer and charge separation are affected in *Arabidopsis thaliana* mutants lacking light-harvesting chlorophyll a/b binding protein Lhcb3. *J Photochem Photobiol B*. **153**, 423–428 (2015).
30. Su, Y. C. *et al.* Small RNA sequencing reveals a role for *sugarcane* miRNAs and their targets in response to sporisorium scitamineum infection. *BMC Genomics*. **18**, 325 (2017).
31. Ahmed, F. *et al.* Comprehensive analysis of small RNA-seq data reveals that combination of miRNA with its isomiRs increase the accuracy of target prediction in *Arabidopsis thaliana*. *RNA Biol*. **11**(11), 1414–1429 (2014).
32. Jia, X. L. *et al.* High-throughput sequencing of small RNAs and anatomical characteristics associated with leaf development in *Celery*. *Scientific Reports*. **5**, 11093 (2015).
33. Xie, K. *et al.* Gradual increase of miR156 regulates temporal expression changes of numerous genes during leaf development in *Rice*. *Plant Physiol*. **158**, 1382–1394 (2012).
34. Mao, Y. F. *et al.* MicroRNA319a-targeted *Brassica rapa* ssp. *pekinensis* TCP genes modulate head shape in *Chinese Cabbage* by differential cell division arrest in leaf regions. *Plant Physiol*. **164**, 710–720 (2014).
35. Wong, C. E. *et al.* MicroRNAs in the shoot apical meristem of soybean. *J Exp Bot*. **62**, 2495–2506 (2011).
36. Wang, L. *et al.* MiR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in *Arabidopsis*. *J Exp Bot*. **62**, 761–773 (2011).
37. Rodriguez, R. E. *et al.* Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. *Development*. **137**, 103–112 (2010).
38. Kuo, H. F. & Chiou, T. J. The role of microRNAs in phosphorus deficiency signaling. *Plant Physiol*. **156**, 1016–1024 (2011).
39. María, J. G., Francisco, J. R., Carlos, L., Esteban, A. & Rafael, P. V. Focus on Ethylene: ethylene and the regulation of physiological and morphological responses to nutrient deficiencies. *Plant Physiol*. **169**, 51–60 (2015).
40. Tao, S. *et al.* The THO/TREX complex active in miRNA biogenesis negatively regulates root-Associated acid phosphatase activity induced by phosphate starvation. *Plant Physiol*. **171**(4), 2841–2853 (2016).
41. Millar, A. A. & Gubler, F. The *Arabidopsis* GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell*. **17**, 705–721 (2005).
42. Alonso-Peral, M. M. *et al.* The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in *Arabidopsis*. *Plant Physiol*. **154**, 757–771 (2010).
43. Achard, P., Herr, A., Baulcombe, D. C. & Harberd, N. P. Modulation of floral development by a gibberellin-regulated microRNA. *Development*. **131**, 3357–3365 (2004).
44. Palatnik, J. F. *et al.* Control of leaf morphogenesis by microRNAs. *Nature*. **425**, 257–263 (2003).
45. Li, Y., Li, C. Q., Ding, G. H. & Jin, Y. X. Evolution of MIR159/319 microRNA genes and their post-transcriptional regulatory link to siRNA pathways. *BMC Evol Biol*. **11**, 122 (2011).
46. Nicaise, V., Roux, M. & Zipfel, C. Focus Issue on Plant Interactions with Bacterial Pathogens: recent advances in PAMP-Triggered immunity against bacteria: pattern recognition receptors watch over and raise the alarm. *Plant Physiol*. **150**, 1638–1647 (2009).
47. Yang, J. *et al.* Auxin response factor 17 is essential for pollen wall pattern formation in *Arabidopsis*. *Plant Physiol*. **162**, 720–731 (2013).
48. Liu, Q. *et al.* Complementation of hyponastic leaves I by double-strand RNA-binding domains of Dicer-like 1 in nuclear dicing bodies. *Plant Physiol*. **163**, 108–117 (2013).
49. Kim, M. *et al.* Activation of the programmed cell death pathway by inhibition of proteasome function in plants. *J Biol Chem*. **278**, 19406–19415 (2003).
50. Sieber, P., Wellmer, F., Gheyselinck, J., Riechmann, J. L. & Meyerowitz, E. M. Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. *Development*. **134**, 1051–1060 (2007).
51. Guo, H. S., Xie, Q., Fei, J. F. & Chua, N. H. MicroRNA directs mRNA cleavage of the transcription factor *NAC1* to down regulate auxin signals for *Arabidopsis* lateral root development. *Plant Cell*. **17**, 1376–1386 (2005).
52. Williams, L., Grigg, S. P., Xie, M., Christensen, S. & Fletcher, J. C. Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. *Development*. **132**, 3657–3668 (2005).
53. Park, H. C. *et al.* Pathogen- and NaCl-induced expression of the SCaM-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. *Plant Physiol*. **135**, 2150–2161 (2004).
54. Shahab, S. *et al.* The effects of microRNA transfections on global patterns of gene expression in ovarian cancer cells are functionally coordinated. *BMC Medical Genomics*. **5**, 33 (2012).
55. Taguchi, Y. H. Possible miRNA coregulation of target genes in brain regions by both differential miRNA expression and miRNA-targeting-sepecific promoter methylation. *ICIC, CCIS*. **375**, 225–230 (2013).
56. Vasudevan, S., Tong, Y. & Steitz, J. A. Switching from repression to activation: microRNAs can up-regulate translation. *Science*. **318**(5858), 1931–1934 (2007).

Acknowledgements

This work was supported by the Natural Science Foundation of China (31570698; 31770743; 31300585). We are thankful to Sichuan Agricultural University, China for providing support in the present investigation.

Author Contributions

J.M. and Y.Y.X. conceived and designed the experiments. Y.Y.X., X.L., R.X.L., S.M.Y. and Z.L. performed the experiments. Y.Y.X., J.M., Y.H.H., F.X.J., X.L., L.X.S. and Z.H. analyzed the data. J.M. contributed reagents/materials/analysis tools. X.Y.Y. and J.M. wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-20261-z>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018